Paparella: Volume I: Basic Sciences and Related Principles

Section 2: Physiology

Part 1: Ear

Chapter 9: Psychophysical Limits of Auditory Performance: Implications for the Diagnosis of Hearing Loss

Donald Henderson, John Mills, Vitorio Colletti

Normal hearing is often taken for granted, and the remarkable acoustic analyzing capabilities of the ear are typically unrecognized or unappreciated. If we pause to study how well the normal human auditory system works, we find that it can detect sound waves with displacements about the diameter of a hydrogen molecule, and can accurately encode sound pressure changes over a range of 10.000.000 to 1. The auditory system has the capability of detecting a single discordant violin while the whole orchestra is playing, and we can localize the source of sound on the basis of a very small (<0.0001 sec) time difference between the two ears.

In modern life, the impressive auditory capabilities of humans are trained and used in spoken language. Often the first hint that a person has a defect in his auditory capability is the awareness of poorer speech perception, particularly when speech takes place in a noisy background. As a first approach to understanding these complains, the audiologist or otologist tests the patient's sensitivity (through an audiogram). Although this may be an adequate diagnostic test for some pathologies, the loss of speech perception ability is more likely to be the consequence of a much larger set of deficits in basic auditory processing than are routinely assessed. Thus, this chapter reviews the sensory capabilities of the normal auditory system as a reference for understanding the auditory capabilities of patients with hearing impairment and their deficits in speech perception. Before discussing these sensory capabilities, it is important to understand how auditory sensations can be studied objectively.

Psychophysics: The Measurement of Sensation

Our sensations are basically private, subjective, and, until recently, not included in the realm of science. In the late 1800s, Gustave Fechner (1860) developed a set of experimental techniques that made the scientific study of sensations possible. Fechner's goal was to bridge the objective physical world with subjective mental experience. Fechner's techniques required that the stimulus be specified precisely (for audition: intensity, frequency, and duration), then the subject could be asked one of three types of questions: (1) how much of the stimulus is needed to be detected, i.e., absolute threshold; (2) how small a change in the stimulus is just detectable, i.e., difference threshold; and (3) how large a change is required to either halve or double the sensation, i.e., sensory scaling. His approach has come to be known as the science of psychophysics. Psychoacoustics is a branch of psychophysics that strives to relate the physical dimensions of auditory stimulation, ie the intensity or frequency of a sound, to its resultant psychological sensation (ie the loudness or pitch of the sound). Thus, even though we cannot share the sensory experiences of another person by using the techniques of psychophysics, we can objectively measure and scale our sensory response to stimulation and compare them with other listeners' responses. Many of the diagnostic audiologic procedures are psychophysical methods that have been adapted for clinical use.

A complete description of the psychophysical methods in audiology is beyond the scope of this chapter. What is important to note, however, is that the psychophysical methods can provide an objective, reliable measurement of an individual's absolute sensitivity, ability to discriminate changes in either the intensity or frequency of a sound and growth of loudness, and ability to localize sound. These are fundamental indices of auditory competence. We shall see that there is not a simple correspondence between the parameters of a sound and the resultant sensation, but rather our sensations are the product of an interaction of the combination of the frequency, intensity, and duration of the auditory stimulus. The following section describes our auditory sensory abilities and provides a perspective for the evaluation of patients with different types of hearing loss.

Absolute Sensitivity

The audiogram is the most common index of a patient's auditory function. Superficially, it is a relatively simple measure. A patient is isolated in a quiet room and the level of an auditory stimulus is controlled by either the clinician or the individual with the aim of determining the level of the stimulus that is just detectable. As we will see, this "quiet threshold" is the product of a number of factors and is often not a simple measure.

Sivian and White (1933) tested quiet thresholds or absolute sensitivity in a number of normal hearing young adults. The results of their investigation are shown as audibility curves in Figure 1 (Dodson and King, 1952). The absolute sensitivity over the range of 20 to 20,000 Hz varies with the frequency of the stimulus, with the best thresholds being in the range of 1000 to 3000 Hz. At the most sensitive frequencies, the stimulus level is less than 0 dB SPL. The decibel unit refers to either the intensity or the pressure of sound. Since it is difficult to measure intensity, the conventional measure of sound is pressure and dB pressure = 20 log(P1/Pr), where P1 is the pressure being measured and Pr is the reference. The typical physical unit of sound pressure is dB SPL, which refers to the use of 20u Pascals (Pa) per cm^2 as a reference (Pr); thus, a threshold value of 20 dB SPL implies the sound pressure at threshold was 200 uPa (20 log(200/20uPa) = 20 dB SPL). This does not mean the absence of sound, but rather the sound pressure at threshold is equal to the reference pressure of 20 uPa cm^2. Here is the first hint of the remarkable sensitivity of the ear, because at 0 dB SPL the ear is detecting displacements of the tympanic membrane that are smaller than the diameter of a hydrogen molecule.

In Figure 1, the lower curve refers to minimum audible field (abbreviated as MAF) and the data are from subjects listening with one ear to sounds coming from a speaker. Threshold is defined as sound pressure measured at the center of the listener's head. The curve labeled MAPC (minimum audible pressure) refers to quiet threshold when listeners are using headphones and the sound pressure at threshold is estimated by using a calibration coupler. The 3 to 10 dB difference between the two curves has multiple causes. The use of headphones for the MAP measurements may add to physiologic noise at low frequencies. The MAP thresholds are only an estimate of what the sound pressure should be at the ear, based on measurements in a supposedly equivalent coupler. However, there is some questions about the accuracy of the calibration procedures.

We should remember that the data from Sivian and White are the average quiet thresholds. Clinicians are usually interested in a particular patient's hearing sensitivity relative to a norm. In the United States, the American Standards Institute 536-1969 defines the average sensitivity as 0 dB hearing level (HL), and for each audiometric frequency at O dB HL, there is a standard physical reference sound pressure expressed as dB SPL. Thus, clinical measures of hearing sensitivity, the audiogram, are reported as deviations from the normal value and are expressed as decibels of HL. Figure 2 shows a typical audiogram and the American National Standards Institute's values for HL at the common frequencies of the audiometer. Again, the reader should be aware that 0 dB HL does not refer to the absence of sound but rather 0 dB HL refers to the standard sound pressure at the different audiometric frequencies listed in Figure 2. Thus the patient's HL at 3000 Hz is 30 dB HL, but in physical units the threshold is 40 dB SPL (30 dB HL, added to the 10 dB SPL for the American National Standards Institute standard at 3 kHz).

The arch shape of the audibility curve seen in Figure 1 may be explained on the basis of a series of mechanical transformations that occur in the outer, the middle, and the inner ear. Dallos (1973) has shown how the frequency of the stimulus is processed and transformed as it proceeds from the external meatus to the basilar membrane. The set of transformations are difficult to acquire for humans, but they have been carefully measured for the cat. Figure 2 A shows how the outer ear and external meatus boost the amplitude of sounds in the midfrequency range. The external meatus acts like a quarter wave resonator, and the resonant frequency (f0) is determined by the length of the meatus with the following formula:

f0 = (speed of sound in air) / 4 x length of meatus).

For the human ear with a 2.5-cm meatus, the peak of resonance is approximately 3.4 kHz. Thus the sensitive midfrequency range of 1500 to 4000 Hz is probably the consequence of the amplification of sound in this range by the resonant characteristics of the meatus.

The poorer sensitivity in the high-frequency region of the audiogram is partially the consequence of the attenuation of the middle ear. Figure 3B shows the transmission of sound through the middle ear as a function of frequency. There is a broad resonance at 1500 Hz; above 1500 Hz, the middle ear attenuates sounds at the rate of approximately 6 dB for each octave increase in frequency. The attenuation above 1500 Hz comes about because of the increased importance of the mass of the middle ear structure for increasing the impedance and attenuation of higher frequency sounds. The external meatus and the middle ear collectively act as an impedance matching device, thereby allowing a more effective transition of airborne sound to vibration in the fluid movement of the cochlea.

Figure 3 C shows the influence of the helicotrema on low-frequency sounds. High-frequency sounds cause a traveling wave that is localized to the base of the cochlea; low-frequency sounds produce a traveling wave that courses through the whole length of the basilar membrane. (See Chap. 8 in this volume). Traveling waves that move to the apical region of the cochlea actually cause fluid movement through the helicotrema (Békésy, 1960). Thus, to some degree, poorer low-frequency hearing is a consequence of impedance provided by the restriction of fluid flow at the helicotrema.

Figure 3 D shows how the combination of transformations of the external meatus, the middle ear, and the helicotrema leads to an audibility curve that reasonably approximates the actual measured audibility curve. Thus, the relative sensitive of the auditory system is primarily the consequence of mechanical-acoustical factors rather than limiting factors inherent in the cochlear hair cells or eight nerve fibers.

The sensitivity of the auditory system is also determined by the duration of the stimulus (Hughes, 1946; and Watson and Gengel, 1969). Sounds of short duration are more difficult to detect than sounds of long duration because they require greater intensity at threshold. Figure 4 shows the increase in threshold as the duration of the signal is decreased from 500 to 10 msec. At longer durations (> 200 msec), the detectability does not appreciably improve, suggesting that the auditory system has an integration of time of approximately 200 msec. The normal integration of acoustic power appears to be a reflection of general sensory processing, because both the visual and the somatosensory systems show integrative processes that are comparable to the auditory system.

The phenomenon of temporal integration has interesting clinical applications (Wright, 1978). Patients with conductive hearing loss have normal integration time, whereas patients with sensorineural hearing loss show reduced temporal integration, ie, there is less than a 10-dB improvement in sensitivity when the test tone is increased from 20 to 200 msec. Limited data has shown that patients with tumors of the eight nerve and central auditory lesions can have much greater difficulty in detecting short duration signals (Olsen, 1974). Although the variation in threshold with the duration of the test signal is a reliable phenomenon, it is of limited clinical value because the range of temporal integration (ie 10 dB) is small compared with the normal measurement error (+ 5 dB) in clinical audiology.

Quiet threshold, probably the simplest of psychoacoustic tests, can be rather complicated. For a given frequency, sensitivity is dependent on a series of mechanical transformations in the outer, the middle, and the inner ear as well as on the duration of the stimulus. Although not discussed here, sensitivity is also influenced by instructions given to the listener.

Loudness

We intuitively equate loudness with the intensity of the stimulating sound. Although this is approximately correct, psychophysical techniques reveal that loudness is also a function of the frequency, band width, and duration of the stimulating sound.

Loudness is a psychophysical dimension of sound and subjective approaches are required to measure or describe the loudness of a given stimulus. Two general approaches to measuring loudness have been developed. One approach, called loudness balancing, requires subjects to match the loudness of a test tone to a second, reference tone. The other approach too measure loudness, called magnitude estimation, requires subjects either to directly scale the loudness value of sound or to physically adjust a sound to achieve a change in sensation, ie, a doubling and halving in loudness.

The phon scale is an example of a balancing procedure for estimating loudness. The reference of standard sound is a 40 dB SPL pure tone of 1000 Hz whose loudness is defined as 40 phons. Other frequencies are matched in loudness to 1000-Hz tones. Thus, the loudness of a given tone is expressed in terms of phons of a matching 1000-Hz tone. For example, a 250-Hz tone has a 40-phon loudness when it has approximately 48 dB SPL (Robinson and Dodson, 1956). Figure 5 shows the family of equal loudness contours for combinations of frequency and SPL that are found on a given contour. Each tone has the same loudness and is expressed in terms of phons or the SPL of the 1000-Hz tone at that contour. The equal loudness curves reveal some interesting features of loudness. First, for a given sound pressure, the loudness varies with frequency, ie, midfrequencies are the loudest and the low and high frequencies are less loud. Second, the differences in loudness between frequencies becomes less prevalent at higher intensities, ie, the equal loudness curves become flatter. Third, the dynamic range for loudness is smaller for low- and high-frequency sounds and greater for midrange sounds. The equal loudness curves reveal the frequency contribution to loudness, but the curves do not directly show how loudness grows with intensity. To appreciate the relationship between intensity and loudness, we need to apply the other psychophysical procedure, ie, magnitude estimation scales.

Another approach to estimating the magnitude of sound is to ask a subject to adjust a given sound so that it is either half or double the loudness of a standard tone. When such manipulations are done over a range of intensities, we find that for every 10-dB increase in intensity, there is an approximate doubling of loudness of sound (Stevens and Newman, 1936).

An alternative approach to scaling sensation was developed by S. S. Stevens (1955, 1961) and is called the magnitude estimation. A subject is presented with a test sound and asked to arbitrarily assign a numerical value (x) to the sound. Then a series of sounds are presented, and the subject is asked to assign each a numerical value that reflects the difference in loudness between the sound (x) and the loudness of each of the sounds in the series. Stevens postulated that sensory magnitude for a whole range of sensory experiences, ie, loudness, brightness, electrical shock, or muscle strain, would be described by the relation,

$$\text{Percept} = KS^n$$

where K = constant, S = physical measure of stimulus, and n = exponent that characterizes the magnitude of a given sensory dimension. For loudness, the exponent, $n = 0.6$, and for comparison, with electrical stimulation, sensation grows with an exponent $n = 3.5$ (Fig. 6). For modalities with exponents larger than one, increasing the stimulus produces a rapid growth of sensation and a limited dynamic range from threshold to saturation (Stevens, 1970). The rapid growth of sensation serves a protection function for certain dangerous stimuli, ie, electrical shock. But for designers of cochlear implants, it is difficult to electrically produce anything close to the normal range of loudness (100 dB) when the electrical stimulus has such a narrow range (10 to 15 dB).

The sensory scale procedures have greater amounts of variability than simple detection or discrimination measurements, but the fact that the scaling data can be reproduced in different laboratories is testimony to their inherent validity. Both approaches show that loudness grows more slowly than the underlying physical stimulus; thus, there is a compression of the range of sensation. Furthermore, loudness also appears to be a function of the duration of the stimulus, but the range of temporal integration above threshold is shorter than temporal integration at threshold (50 to 100 msec versus 200 msec).

The test of an individual's growth of loudness has been used clinically as an indicator of cochlear function. Loudness measurements in a clinical setting have been either monaural or binaural procedures. When an individual has normal hearing in one ear, then a binaural loudness balancing test can be used (Carver, 1978). Tones of the same frequency are presented to each ear, the intensity of the tone in the impaired ear is varied, and the subject is instructed to indicate when the two tones are equally loud. Figure 7 gives a representative sample of typical test results. The solid line is typical of two normal ears, ie, normal thresholds and essentially normal growth of loudness; the dotted line is typical of a conductive loss, ie, the right ear is less sensitive, but with suprathreshold stimulation, loudness grows equally in both ears; the dotted line is typical of patients with sensorineural hearing loss, ie, the right ear is less sensitive, but loudness grows rapidly above threshold and in the impaired ear, 30 dB above threshold loudness is essentially the same as the normal ear. This phenomenon of loudness growth is termed recruitment and is consistent with peripheral sensorineural loss. The presence of recruitment is a factor to be considered when fitting an individual with a hearing aid. In rare cases of central auditory pathology, loudness grows slower than normal and is termed decruitment (Carver, 1978).

Loudness growth, in certain types of patients, can be measured with monaural techniques. Tones can be selected from the normal and the impaired regions of the audiogram and matched for loudness. Thus, the growth of loudness for normal frequencies can be compared with the growth of loudness for frequencies from the impaired range. Monaural loudness techniques are even more difficult than binaural techniques, and their cost in time and inherent variability limits their use in most clinical settings.

Intensity Discrimination

The ability to detect small changes in intensity is an important part of everyday listening. In terms of sensory psychology, the smallest detectable change in intensity is called the difference limen for intensity, or delta I. Since the beginning of experimental psychology, sensory discrimination abilities were described by the Weber fraction: delta S/S = constant.

The Weber fraction follows from our experience that the ability to detect a change in sensory stimulation depends on the base level of stimulation, eg, we can feel the difference between a 10- and 11-gram weight, but we would need a weight to be at least 110 grams to be detected as different from a 100-gram weight. Tests for the determination of delta I require a listener to detect either the increment or decrement of a pure-tone stimulus or to select the louder of two sinusoidal tone bursts. Each of these techniques has limitations, but if done carefully, delta I can be determined for a large range of intensities and frequencies.

Figure 8 is from Jesteadt and associates (1977) and shows the range of delta I across a wide range of listening conditions. Several points are made by this set of experimental results. First, delta I is independent of the frequency of stimulation over the range of 250 Hz to 8000 Hz. Second, delta I is poorest at the lowest levels of stimulation and improves with increased levels. Third, delta I systematically varies between 0.5 and 1.5 dB over an 80-dB range of stimulation. These results have been described as a "near miss" to Weber's law because Weber's law predicts a constant delta I/I, and in actual measurement delta I only changes about 1 dB over a large range of audibility.

The relationship between the growth of loudness and delta I has been used to estimate recruitment in hearing-impaired listeners. It follows that a faster growth of loudness (ie, large change in loudness for small change in intensity) would be accompanied by smaller delta I, thus the presence of recruitment could be predicted by an abnormally small delta I. This was the hypothesis underlying the Small Increment Sensitivity Index (SISI) (Martin, 1978), in which patients are requested to listen for intensity increments to a continuous tone presented at 20 dB above their thresholds. Even though the test has some scientific basis, it has faded from most clinical batteries.

Pitch

As a first approximation, pitch appears to be the psychological quality associated with the frequency of the stimulus. Again, as with loudness, it is difficult to measure and study the perception of pitch. However, using scaling techniques, it is possible to relate pitch to the frequency of the stimulating sound. The unit of the pitch scale is called the mel (Stevens and Volkmann, 1940), and the pitch associated with a 1000-Hz tone at 40 dB SPL is arbitrarily defined as 1000 mels. By instructing the listener to adjust the frequency of the sound so that the pitch is either doubled or halved, the relationship between frequency and pitch was mapped by Stevens and Volkmann. Typical results are given in Figure 9. The auditory system compresses the band width of perception, making the range of pitch (mels) considerably smaller than the range of frequencies that are heard.

The pitch of a sound is also dependent on the duration and intensity of the stimulus. Experiments by Doughty and Garner (1947) showed that listeners needed 5- to 10-msec samples of a sound before they could adequately assign a pitch to the sound. Other experiments by Stevens (1935) showed that the pitch of a sound systematically changes with increases in intensity. For low-frequency sounds (< 1000 Hz), the perceived pitch decreases as the intensity of the sound increases; conversely, the perceived pitch of high-frequency sounds increases with high levels of stimulation, as shown in Figure 10. These are difficult measures to make and these data have not been completely supported in subsequent studies, but it is interesting to note that the shifts in pitch are consistent with how the pattern of traveling wave activity in the cochlea changes with intensity.

Pitch perception has been one of the most studied and debated aspects of auditory theory. The dilemma about how pitch is encoded in the cochlea becomes more interesting when the stimulus is complex rather than a simple sinusoid. Experiments by Schouten (1962) and colleagues frame the dilemma. They presented subjects with a complex tone composed of individual sinusoidal components at 700, 800, 900, and 1000 Hz. Even though there is no energy at 100 Hz, the subjects report the perception of a 100-Hz tone or the common, but missing, fundamental. The results of this and other complementing experiments suggest that pitch, particularly for low frequencies (< 2000 Hz), is dependent on the temporal patterns of the stimulus.

A patient's pitch perception ability is rarely tested in clinical settings probably because of the difficulty in measuring pitch. However, unilateral Ménière's disease (Tonndorf, 1968) or sensorineural loss from noise exposure can produce diplacusis so that the perceived pitch of a given stimulus is different when the same stimulus is delivered to each of the ears. Diplacusis is an unlikely complaint in an audiologic evaluation because patients apparently develop compensatory strategies for deficits in the ability to perceive differences in pitch. The ability to match pitch may have clinical utility for patients with tinnitus. The clinician can have a better perspective on a patient's complaint if the patient can match their tinnitus to a pure tone or narrow-band noise stimulus.

Frequency Discrimination

Although absolute pitch judgments are not that important for daily listening, the ability to make pitch discriminations is quite important, particularly for effective speech perception. The ear can detect minor changes in frequency; however, the actual limits of our frequency discrimination ability is partially determined by the measurement procedure. Shower and Biddulph (1931) frequency discrimination by having observers detect frequency modulation of a carrier sound. A recent study by Wier and associates (1977) used a two-alternative forced choice or a comparison test that required pitch memory. Their results provide the most comprehensive perspective on frequency discrimination as a function of both the frequency and intensity of the stimulus. In Figure 11, delta F is plotted against frequency. It is clear that for low frequencies (< 1000 Hz), delta F is a constant number of cycles, but above 1000 Hz, delta F becomes progressively larger. Also, at low intensities (5 and 10 dB SL), delta F is poorer but remains relatively constant from 20 dB to 80 dB SL.

Frequency discrimination is degraded with sensorineural hearing loss, but the relationship between the size of the hearing loss and the increase in delta F is not clear (Turner and Nelson, 1979). Furthermore, delta F is not part of the clinical battery because of the difficulty in reliably measuring the size of delta F. We will see in the next section how the increase in delta F is probably the consequence of broader tuning in the auditory system.

Masking and Frequency Selectivity

Our discussion of auditory listening abilities has been limited to simple sinusoidal signals without background noise. Although such measures are appropriate for the laboratory setting, realistic listening situations often involve the auditory processing of complex sounds (time-varying, multicomponent sounds) in a noisy background. The background noise can be a deterrent to clear auditory perception or frequency analysis. The elevation in detectability of a sound by the presence of other background sounds is called masking. In a sense, masking is the failure of frequency analysis, and quantitatively, the degree of masking is the amount of elevation in threshold of a signal by the background.

An appreciation of the rules of masking can be gleaned from studying simple masking by pure tones. In the first study by Wegel and Lane (1924), the detectability of pure tones ranging from 400 to 4000 Hz was measured in the presence of a 1200-Hz masking tone. As can be seen in Figure 12, the pattern of masking is quite complicated. When the frequency of the signal is appreciably below the 1200-Hz masker, there is essentially no masking; masking grows as the signal approaches 1200 Hz; when the two tones are close in frequency, the signal becomes easier to detect because the signal and masker are producing acoustic beats; as the signal moves to higher frequencies, there is some decay of masking; when the signal is at the second or third harmonic, detectability improves because the two tones beat again. When the two tones are close in frequency, they move in and out of phase with each other and the envelope of the two tones increases and decreases at a rate equal to the difference in frequency of the two components. The subject detects the change in the envelope of the sound as an acoustic beat.

Thus, the masking patterns reported by Wegel and Lane are complicated by the phenomena of beats and difference tones. In a nonlinear system, when two tones are added together it is possible to produce difference or distortion tones equal in frequency to f1 to f2, f2 to f1, 2f1 to f2, and so on. Some of these distortion tones are stronger than others. If these acoustic parameters are minimized or controlled by using short duration tones, minimal phase locking between the masker tone can be detected, yielding more simple masking patterns. Figure 13 reports examples of psychophysical tuning curves - a measure that is related to the masking patterns of Wegel and Lane. The psychophysical tuning is measured with a probe tone of fixed frequency and low sensation level. A masking tone of varying frequency and intensity is added. The subject has the task of detecting the probe tone as the masker is adjusted in level and at frequencies above and below the probe tone. Presumably, the low-level probe excites a limited population of hair cells and fibers of the eight nerve; thus the shape of the tuning curves in Figure 13 are assumed to be a reflection of the spread of mechanical activity on the basilar membrane.

The psychophysical tuning curves reveal more about the pattern of masking and frequency selectivity. For high-frequency signals (3000 Hz), the masking pattern is asymmetric. The high-frequency leg of the tuning curve is steep, whereas the low-frequency leg is shallower and has an even shallower "tail" 30 to 50 dB above the tip of the tuning curve. For low frequencies (300 Hz), the tuning curve is much shallower and more symmetric (Wightman et al, 1977).

The psychophysical tuning curves are considered to be a reflection of the filtering action of the basilar membrane (see Chap. 8). A comparison of tuning curves of individual fibers of the eight nerve and psychophysical tuning curves is provided in Figure 14 (Salvi et al, 1982). These data come from an experimental chinchilla who had been exposed to noise. Two features are interesting. First, there is a close correspondence between the psychophysical tuning curve and the neural tuning curves. Second, in frequency regions with good hearing, the tuning curves have normal appearance; at frequencies with hearing loss, both the neural and psychophysical tuning curves are much broader. The obvious implication of these results is that individuals with sensorineural hearing loss are not only less sensitive, but the broader tuning of the system is an impediment to resolving signals in a noisy background.

The tone-on-tone masking patterns or tuning curves are reliable phenomena and illustrate some of the ear's frequency-resolving processes. However, in real life, the more typical masking situation is when a signal is masked by a noisy background. When a listener is trying to detect a signal of frequency (x) in a broad-band noise, only the component close in frequency to (x) actually contributes to the masking of the signal. The effective band width of the noise that contributes to the masking is called the critical band width. Fletcher (1950) first reported the critical band width conception in an experiment in which he studied the detection of signals in a noisy background and then systematically narrowed the band width of the noise. For signals of 500 Hz, the limits of a broad-band noise could be moved closer to 500 Hz without a release of masking. Only the components in the noise from approximately 425 Hz to 575 Hz (the critical band) actually contributed to the masking of the signal. The size of the critical band increases systematically above 500 Hz (Fig. 15). The critical band has been measured by using a number of different techniques (Scharf, 1970), and the results show enough consistency to conclude that the critical band is a reflection of how information about the frequency content of a complex signal is organized for processing in the central auditory system.

Masking, over a relatively wide range, is a linear phenomenon, ie, when a signal is at masked threshold, and the masking noise is increased 10 dB, then for the signal to be just detectable again, it needs to be increased another 10 dB. The linearity of masking, as can be seen in Figure 16, covers a wide range of frequencies and intensities (Hawkins and Stevens, 1950).

Temporal Processing

We have seen how the basic sensory qualities of pitch and loudness are primarily determined by the fundamental physical dimensions of sound, ie, frequency and intensity. The temporal pattern of a sound is the third primary physical dimension of a sound. However, the role of the temporal pattern in the production of sensation is not as clear as the effects of frequency and intensity. Earlier in this chapter (see Fig. 4), the effect of signal duration was seen to be important for both loudness and for decreasing quiet threshold. We shall now turn to a discussion of other aspects of temporal pattern, ie, temporal resolution and temporal masking.

The temporal relation between a sequence of sounds can be important for the detection and identification of complex sounds. For example, the detection of a specific segment in a series of sounds can be obscured or masked by sounds just preceding or following it. Figure 17 shows a schematic of temporal masking. On the left side of Figure 17 is backward masking, or when the signal precedes the masker. The actual period of time between the signal and the masker is quite small; thus, backward masking is probably not a factor in much of everyday listening. (Neurologically, backward masking is probably the consequence of a more intense stimulus (the masker) being conducted faster through the auditory system than the lower intensity signal. There, the masker "catches up" with the signal and masks or obscures the sensation normally produced by the less intense signal (Wilson and Carhart, 1971).

For comparison, forward masking is shown on the right side of Figure 17. In contrast, forward masking occurs over a much longer period of time than backward masking and may be a factor in the perception of certain speech sounds (Wilson and Carhart, 1971). Neurologically, forward masking is probably the result of the relative refractory period of the auditory system following the masker stimulus. Thus, it is not surprising that the amount of forward masking increases with progressively more intense signals and noise.

Listeners are often faced with the task of detecting silent intervals for extracting "meaning". For example, the recognition of a group of speech sounds is dependent on the duration of the voice onset time or the silent period between the initial part of the consonant and the onset of voicing. The minimum period of quiet time that can be detected has been named by Green (1971; 1973) the "minimum integration time". Operationally, the minimum integration time can be measured by presenting a listener a noise stimulus with a silent "gap" and measuring the duration of the gap that is just detectable. Over a relatively wide range of band widths and intensities, both humans and chinchillas have similar gap thresholds, with the best resolution occurring for signals with broad bands and intensities 30 dB above threshold (ie optimal conditions nurture gap detection thresholds of 3 msec).

Both gap detection and forward masking are systematically degraded with sensorineural hearing loss. Nelson and Turner (1980) showed that patients with sensorineural hearing loss had greater amounts of forward masking. In a series of experiments in our laboratory (Arehole et al, 1987), chinchillas were given either flat hearing loss ranging from 15 to 60 dB or high-frequency losses of approximately 40 dB. The general trend of the data showed that losses of 30 dB or less do not compromise temporal resolution or forward masking; losses greater than 30 dB degrade temporal resolution and prolong forward masking. Gap detection, in particular, is dependent on the integrity of hearing the high frequency range.

The gap detection and forward masking results obtained from an experimental animal model of sensorineural hearing loss, but they are consistent with human clinical studies. Furthermore, the clinical significance of the relationship between hearing loss and temporal processing is quite interesting. Tyler (1986) has shown that in a psychoacoustically measured battery, gap detection thresholds were correlated with the subject's speech perception performance.

Binaural Hearing. In most animal species other than humans, the ability to localize a sound is as important as the ability to identify a sound. For many of the auditory capabilities that we have just discussed, binaural listening only provides marginal improvement over monaural listening. However, for optimal localizing ability, two ears provide a major advantage by processing the differences in the acoustic signals at each ear.

The classic experiment on auditory localization was performed by Stevens and Newman (1936) on the roof of the psychology department building at Harvard. They moved the experiment out of the lab to avoid the confusing sound reflections found in closed rooms. Their subjects were blindfolded, seated on elevated chairs, and asked to point to a sound source that was moved in an arc at a distance from their head. Figure 18 shows the trend of the results: At low frequencies (up to 900 Hz), the subjects localized the source with minimal error; as the frequency of the source was increased, their performance deteriorated with maximum error at 3000 to 4000 Hz; with further increases in signal frequency, performance improved.

To appreciate the relationship between localization ability and frequency of the stimulus, it is necessary to understand the differences in sound stimulation at the listener's ears. Consider the sound and head interactions when the sound source is directly opposite one ear. At low frequencies, the size of the head is small compared to the wavelength of the sound; consequently, sound "wraps" around the head with minimal attenuation and, thus, the primary acoustic differences between the ears are the time of arrival and the phase differences. Consequently, low-frequency tones are localized on the basis of temporal cues. In the midfrequency range (1000 to 5000 Hz), the head still does not provide much of a barrier to reflect sound and the arrival times between the two ears (approximately 0.6 msec) is long compared with the period of the sound, but the interaural time difference for this midfrequency region is very short relative to that for the low-frequency region. Thus, the listener does not localize these sounds well because there is not a clear intensity difference between the ears and the time difference can be confusing. At high frequencies, the head becomes an "effective barrier" to sound, producing 5- to 15-dB differences or "sound shadows" between the ears. This intensity difference at the ears provides the cues for localization of high-frequency sounds.

The original experiment performed by Stevens and Newman established the importance of time and intensity cues for localization of sounds. If we present listeners with identical sounds played through headphones, then they hear a fused auditory image in the center of their head. With stimulation via headphones, we can independently control the sounds at each ear and can systematically move or lateralize the sound in the perceptual space between the two ears. Thus, the significance of time and intensity cues can be evaluated independently of each other and can be presented in opposition in the study of the trading relationship between time and intensity (msec/dB), ie, if the left ear is leading by 0.1/msec, how many decibels more intense does the right ear need to be to produce a fused image in the midline?

It is generally accepted that the beginning neural processing for the localization of sounds occurs in the hindbrain, probably in the olivary complex (Jeffress, 1972). Thus it is not surprising that localization or lateralization tasks may be useful indices of the integrity of the brain stem. Bosatra and Russolo (1975) tested the lateralization abilities of normal listeners and patients with a variety of peripheral and central auditory problems. They asked their subjects to decide whether they perceived the stimulus in the midline or lateralized to one side of their head and measured the minimal delta I. An interesting group of patients with brain stem disorders showed much larger delta I, even though they had normal pure-tone audiograms and normal auditory pattern discrimination. Creative exploration of the binaural system may eventually lead to a useful strategy for evaluation of the integrity of components of the central auditory system and may eventually be part of a larger test battery including auditory brain stem response, positron-emission tomography, computerized axial tomography, and magnetic resonance imaging.